

REVISED -  
see 12/2/93 MEMO

IRM S/F

MEMORANDUM FOR: John C. Hoyle  
Assistant Secretary of the Commission

FROM: Arnold E. Levin  
Licensing Support System Administrator

SUBJECT: TECHNICAL WORKING GROUP REPORT

NOV 24 1993

In response to requests raised at the October 5-6, 1993, meeting of the Licensing Support System Advisory Review Panel (LSSARP), I am providing a copy of the Licensing Support System Technical Working Group's report to the Chairman. This report can be provided to the LSSARP members.

It should be noted that the Working Group relied heavily on cost figures and technical information developed for the Department of Energy as part of a 1989 study by Science Applications International Corporation (SAIC). Significant technological advances have occurred in this area since the SAIC study was completed. The original SAIC design also was driven, in part, by compliance with the LSS rule (10 CFR 2, Subpart J). For these reasons, cost figures presented in the Working Group's report differ somewhat from more recent figures presented to the Commission (SECY-93-107).

It should also be noted that some of the documents cited in the Technical Working Group's report may not yet be available to the public.

If you have questions regarding this information, please contact me at 492-7649.

Original Signed By:

Arnold E. Levin  
Licensing Support System Administrator

Enclosure:  
As stated

cc: H. Thompson, Jr., DEDS

Distribution:  
w/o enclosure: G. Cranford  
J. Shields  
D. Drapkin  
E. Shelburne

DOCUMENT NAME: P:\LSSARP.INF

To receive a copy of this document, indicate in the box: "C" = Copy without enclosures "E" = Copy with enclosures "N" = No copy

OFFICE	IRM/DD								
NAME	AELevin:jcv								
DATE	11/24/93								

OFFICIAL RECORD COPY



UNITED STATES  
NUCLEAR REGULATORY COMMISSION  
WASHINGTON, D. C. 20555

NOV 24 1993

MEMORANDUM FOR: John C. Hoyle  
Assistant Secretary of the Commission

FROM: Arnold E. Levin  
Licensing Support System Administrator

SUBJECT: TECHNICAL WORKING GROUP REPORT

In response to requests raised at the October 5-6, 1993, meeting of the Licensing Support System Advisory Review Panel (LSSARP), I am providing a copy of the Licensing Support System Technical Working Group's report to the Chairman. This report can be provided to the LSSARP members.

It should be noted that the Working Group relied heavily on cost figures and technical information developed for the Department of Energy as part of a 1989 study by Science Applications International Corporation (SAIC). Significant technological advances have occurred in this area since the SAIC study was completed. The original SAIC design also was driven, in part, by compliance with the LSS rule (10 CFR 2, Subpart J). For these reasons, cost figures presented in the Working Group's report differ somewhat from more recent figures presented to the Commission (SECY-93-107).

It should also be noted that some of the documents cited in the Technical Working Group's report may not yet be available to the public.

If you have questions regarding this information, please contact me at 492-7649.

A handwritten signature in cursive script, reading "Arnold E. Levin", is positioned above the printed name.

Arnold E. Levin  
Licensing Support System Administrator

Enclosure:  
As stated

cc: H. Thompson, Jr., DEDS

**LICENSING SUPPORT SYSTEM  
WORKING GROUP**

**PRELIMINARY REPORT**

**February 21, 1992**

**NOTE: The Working Group recommendations may not reflect (see Attachment 2) accurately the DOE opinion on each of the proposed alternatives.**

## LICENSING SUPPORT SYSTEM

### Introduction

In an August 29, 1991 letter (Ref. 1) to the Nuclear Regulatory Commission (NRC), the Department of Energy (DOE) proposed a joint DOE/NRC reexamination of the current design, development, implementation and operational aspects of the Licensing Support System (LSS). The LSS is a conceptual electronic information management system proposed for use in the licensing proceedings for a national high-level radioactive waste repository.

The DOE proposal recognized the many evolutionary changes in the civilian high-level radioactive waste program, a revision in the schedule for repository license application, and the concept of the LSS as part of a comprehensive, integrated information management system used to support the development and review of that affiliation.

In response (Ref. 2), the NRC agreed with the DOE proposal to reexamine the LSS design "with the objective of developing a system that is responsive to the needs for improved information management and retrieval, yet more cost-effective than that initially proposed." The response also recognized the regulatory oversight and auditing mission of the NRC, and the opportunity to cooperatively reassess a large (multi-year) system development initiative to ensure that it is both technically sound and cost effective for all participants in the licensing process.

To develop this cooperative analysis, a technical working group was formed (Ref. 3) to reexamine the LSS design. The Working Group, which included representatives from NRC and the DOE's Office of Civilian Radioactive Waste Management (OCRWM), was comprised of the individuals listed below.

Gerald Cranford (Chairman)	NRC/IRM
Dave Drapkin	NRC/IRM
Dan Graser/Barbara Cerny	DOE/OCRWM
Jim Shields	NRC/LSSA
Donna Sitterson/Ray Godman	TRW (DOE Contractor)
John Voglewede	NRC/IRM

After the Working Group was formed, the Secretary of the Commission provided a representative (John Hoyle/Andy Bates). Other individuals from both DOE and NRC have also participated in the discussions of the Group. Representatives from major NRC Offices (ACNW, ADM, ASLBP, LSSA, OGC, NMSS, RES, SECY) have presented their views on the LSS to the Working Group for consideration.

### Technical Review

The Working Group's technical review of the LSS included a number of documents prepared either by the Department of Energy (Refs. 4-5) or for DOE by Science Applications International Corporation (SAIC, Refs. 6-11). The documents were part of a larger collection of information

that describe the proposed design, design considerations, and costs of the LSS.

The Working Group also reviewed documents prepared by the NRC Office of the Licensing Support System Administrator (LSSA) for the Commission. These documents discuss proposed LSSA program and budget responsibility (Ref. 12), Commission responses (Ref. 13), and other Commission briefing materials (Refs. 14-15). In addition, the Working Group reviewed correspondence between the LSSA and other NRC organizational components (Refs. 16-20). Under the terms of the LSS rule (10 CFR Part 2, Subpart J, Refs. 21-22), DOE is responsible for design and development of the LSS and NRC is responsible for system management, including operation and maintenance.

### LSS Working Group Charter

In responding to the DOE/NRC agreement, the Working Group has adopted the following guidelines:

- Examine the original SAIC-proposed LSS design to determine if it remains technically sound, if it conforms to current technology, and if the design supports cost reduction measures.
- Examine INFOSTREAMS to determine if all or parts of that system can be applied to handle the LSS requirements. INFOSTREAMS is a system currently being developed by DOE/OCRWM to automate the collection, storage, and retrieval of records generated and used within the OCRWM organization.
- Evaluate various cost-reduction/reallocation measures to determine if they are technically acceptable and feasible.
- Consult NRC office representatives to determine whether enactment of the cost reduction measures would compromise the functionality of the LSS to the extent it would not meet their needs.
- Make recommendations to the Commission that would minimize development and operational costs of the LSS.

### The SAIC Design

The Working Group has examined the preliminary design for the LSS proposed by SAIC. SAIC's preliminary design was developed under a multi-year DOE contract to be responsive to the LSS rule (10 CFR 2, Subpart J). SAIC conducted an extensive requirements study, data scoping, conceptual design, feasibility analyses and developed and operated a prototype system. This prototype included processing, cataloging, scanning and loading 100,000 pages of LSS documents into a text and image database that typical users then accessed in a controlled environment. Based on the results of these tests and analyses, SAIC developed system-level requirements and preliminary design documents in conjunction with OCRWM and NRC technical staff.

The Working Group found that the original SAIC design is thorough and consistent with automatic data processing system design methods and

practice. The original SAIC design reflects specifications resulting from compliance with the LSS rule.

Although the SAIC study was completed in 1989, this does not seriously compromise the LSS design. DOE has provided evidence (Refs. 23-24) of its continuing efforts to stay abreast of the technology involved in the LSS program; particularly advances in optical character recognition, user interface design, and the operation of large full-text databases; and will continue to monitor this environment for cost saving techniques resulting from advances in technology.

Processing data structures such as text, voice and images is currently an issue of great interest among software and hardware developers. The Working Group believes that cost reductions will be achieved in the automation of text and image processing as the development of the LSS proceeds.

- The Working Group concludes that, given the anticipated size of the relevant document collection and the constraints of the LSS rule, the SAIC-design is appropriate and the system designed by SAIC, is not loaded with features that can be deleted to achieve measurable cost savings without significant reduction of the functionality and coverage of the system.
- Based on the Working Group's review of the SAIC design, and the continuing efforts of the DOE to monitor this technology, it was concluded that an examination of other government and industry offerings would duplicate previous efforts and would not result in the accumulation of significant new information.
- The Working Group believes that technological breakthroughs in the automation of text and image processing will result in lower costs over the life cycle of the LSS. For example, the projected cost of the original SAIC model was reduced by \$14 million with the introduction of differential (multiple) optical character recognition methods (Ref. 11).

#### INFOSTREAMS

The DOE Office of Civilian Radioactive Waste Management has identified the need to automate the collection, storage and retrieval of records generated and used within the OCRWM organization. A subset of these records would be identified for inclusion in the LSS. DOE has contracted with TRW Environmental Safety Systems (TRW) to provide such an automated system. A description of the proposed INFORMATION Storage/REtrieval/Access Management System (INFOSTREAMS) concept has been presented to both the Commission (Ref. 25) and to the LSS Working Group (Ref. 26). DOE has proposed that INFOSTREAMS be used instead of the LSS to convert its LSS documents into the electronic form needed for loading into the LSS database. The LSSA has suggested that software and procedures developed for INFOSTREAMS be examined for application to LSS and NRC needs (Ref. 27).

The INFOSTREAMS design, based on DOE/OCRWM document management system requirements, has some features in common with the SAIC design for the

LSS. INFOSTREAMS will be capable of serving as the capture station for DOE's share of the proposed LSS database (80 - 86 percent). However, the capabilities of the two systems (INFOSTREAMS and LSS) are not identical. INFOSTREAMS will provide a bibliographic index to documents tracked by OCRWM. Digital images will be provided for some of those documents. A smaller subset will be available in full text. The collection of documents in full text will change over time as new documents are added and older documents are retired (i.e., archived and no longer available on-line). This design requirement may be contrasted with the LSS design in which (according to the LSS rule) all information (20 million pages by 2001; 32 million by 2004) will be indexed and made available in full text and on-line image throughout the life of the system.

- The Working Group recognizes a substantial similarity among the document processing and retrieval needs of NRC's high level waste program, OCRWM and the potential parties to the repository licensing proceeding.
- INFOSTREAMS has been designed to meet DOE/OCRWM internal document management requirements and provide input for DOE documents to the LSS. The extent to which INFOSTREAMS hardware and software designs are applicable to the LSS (or to NRC document management requirements) has not yet been determined although INFOSTREAMS (as planned) does provide many of the functions required by the LSS and NRC document management systems.

#### NRC/DOE Office Views

The Working Group invited several major NRC Offices (ACNW, ADM, ASLBP, LSSA, OGC, NMSS, RES, SECY) and DOE/OCRWM to present their views on the LSS. Although the offices interviewed indicated that some features of the LSS were more important to them than others and that they had less need for some features than other offices might, no office recommended that any major feature of the LSS e.g., full text search, image retrieval, bibliographic descriptors, pre-hearing availability of the database, electronic mail, and so forth could be eliminated without undesirable impact on the pre-licensing review or the proceeding itself.

While affirming the expected usefulness of the LSS in the licensing review process, NMSS did point out that the relevant information required by the technical staff may be considerably less than the total backlog of potentially discoverable documents currently planned for inclusion in the LSS database.

NMSS, ACNW, and RES believe that the ability to search the full text of technically relevant documents will be useful. NMSS and ACNW asserted that access to on-line digital images, or other readily available forms of graphic information are important attributes and that eliminating this capability from the LSS would seriously impair its use in the technical review process.

NMSS also noted that if the LSS is not available to the technical staff until late in the program, then it will not be of much use in

the prelicense application phases. They indicated that, for the technical review process, it would be beneficial to implement a fully functional LSS as soon as possible, even if it were necessary to limit the number of documents in the database in order to do so. It was suggested that such a limitation could be accomplished by restricting the processing of backlogged documents to some more recent start date (e.g., the publication date of the Site Characterization Plan). NMSS also noted that although an early license application is not presently expected, it would be prudent to have a contingency plan for such an event (Ref. 30).

The Office of the Secretary suggested a very different course of action -- that consideration be given to delaying the implementation of the LSS as long as possible because of anticipated funding constraints and technological uncertainties, and limiting the LSS when implemented to a "core" database. Non-core documents would be made available to parties by placing them in the public domain well in advance of the hearing as they are generated. Copies could be given to each party if requested. The parties could review indexes of these documents periodically and determine whether certain ones should become "core" documents in the LSS. The LSS would not have to be brought on-line until needed for searching core documents. Implementing this suggestion would offset the total document burden of the LSS by placing information into the hands of interested parties through local and agency Public Document Rooms, thereby reducing the number of documents in the LSS and lowering LSS costs.

OGC and ASLBP noted that the LSS feature most critical to them was full text access to docket material, but that other participants would likely have broader needs. Both offices also voiced the opinion that non-federal LSS participants would react negatively to any proposal by NRC to limit the scope or availability of the LSS.

ASLBP reaffirmed their belief that it would not be possible to meet the Congressionally-mandated hearing schedule without an LSS of the SAIC-type design, and that making relevant information equally available to all participants will shorten the hearing process.

In an October 29, 1991 follow-up memorandum (Ref. 31), ASLBP recommended that: (1) IRM conduct a cost benefit study to verify ASLBP's assertion that the LSS original design would effect enormous savings; and (2) future IRM or NRC investigation and study of the LSS should be conducted jointly with representatives of potential parties because the configuration of the LSS will be so closely intertwined with the procedural and due process.

DOE expects that its staff and contractors will use the LSS to access the documents of other LSS participants as well as some DOE documents (because DOE does not plan to keep all the documents that it processes for inclusion in the LSS permanently on-line in INFOSTREAMS). DOE does not believe any major function of the SAIC design should be eliminated.

DOE has offered to make INFOSTREAMS technology available to NRC if requested. DOE has noted, however, that INFOSTREAMS is being designed



to meet OCRWM's internal needs and is not being functionally designed to meet the LSS needs nor is it scaled to handle the LSS' size requirements. INFOSTREAMS is being designed to use DOE's existing Digital Equipment Corporation VAX/VMS hardware and BASISPLUS software. DOE note that additional customization would be needed to re-develop relevancy ranking algorithms embedded in INFOSTREAMS' content/relevancy based retrieval software. Furthermore, the Department of Energy's contractor (TRW) has emphasized that scaling requirements are a serious concern for migration of the INFOSTREAMS software to the LSS and DOE agrees with this assessment. Only selected information will be made available in image and full-text format to DOE users at any given point in time. Unlike the LSS, which will maintain all relevant high level master repository licensing material on-line for access, INFOSTREAMS will maintain only that portion of information currently being evaluated and will replace that information when a particular need has been met. Moreover, DOE has no plans for permitting outside access to INFOSTREAMS as a substitute means of accessing the text or images of LSS documents.

All the organizations consulted were unanimous in the opinion that any significant deviation from the functional requirements for the LSS would be viewed negatively by the non-federal participants in the rulemaking process. The LSS rule would require renegotiation with uncertain results.

#### Potential Measures to Reduce or Reallocate LSS Costs

Based upon the working group's interest to identify opportunities for possible cost reductions, the Working Group examined a number of options affecting the cost of the LSS program. Following is a list of these options.

1. Capture DOE's LSS material using INFOSTREAMS versus the LSS
2. Adapt INFOSTREAMS search and retrieval capability to the LSS
3. Remove LSS telephone connect charges from the LSS Budget
4. Remove local LSS costs (workstations/other) from the LSS Budget
5. Remove costs to "pre-process" material for input to the LSS
6. Eliminate one of two training/user support locations
7. Use existing infrastructures (rather than LSS) to provide paper copies of LSS documents on demand
8. Minimize non-DOE document capture costs:
  - a. DOE capture all non-DOE LSS material; or
  - b. LSS capture non-DOE LSS material using software adapted from INFOSTREAMS
9. No on-line images/limited on-line images

- a. Digital images of all pages stored off line in LSS; paper copies of requested pages mailed to individual users; or
  - b. No digital images (off line or on line) in LSS; microform images of all pages sent to key user locations; or
  - c. Digital images of "graphic" pages only stored on line in LSS; digital images of text pages sent to key user locations
10. Eliminate some documents from LSS database and place them in NRC/DOE public document rooms in a timely manner
- a. Establish a cutoff date unless to be relied upon; or
  - b. Include only "core documents" in the database
11. Reduce early availability of the LSS database
- a. Limit use and size of the database from early 1996 until early 1998; or
  - b. Don't make database available until early 1998

#### Impact of Working Group Recommendations

The chart on the next page (Figure 1) summarizes the LSS concept/design alternatives considered by the Working Group and provides additional information regarding cost reduction, the benefit reduction (if any) to potential LSS users, the impact on the LSS rule and shows whether or not the alternatives are recommended by the Working Group for adoption.

To assess the overall cost impact of its recommendations, the Working Group used \$193 million as a cost baseline. This baseline was prepared in June 1990 by the LSS Administrator relying almost exclusively on the SAIC revised cost analysis (Ref. 11). It encompasses LSS program costs from FY 1987 through FY 2000 irrespective of who's budget they might fall into. To arrive at a budget for the future development and operational costs of the LSS, the Working Group has reduced the \$193 million estimate as shown below:

<b>\$193.0M</b>	<b>Baseline program costs through FY 2000</b>
<b>-73.6</b>	<b>Move DOE capture from the LSS to the DOE budget -- DOE recommends that its LSS material be processed through INFOTREAMS; DOE would perform most of this processing for its own internal purposes regardless of LSS needs</b>
<b>-9.3</b>	<b>Savings -- Working Group recommendations to take advantage of software developed through INFOTREAMS and to delay availability of full LSS</b>
<b>-9.0</b>	<b>Move non-DOE capture from LSS to DOE budget (INFOTREAMS)</b>

-7.9      Move telecommunication charges, LSS workstations, and  
document pre-processing to LSS participant budgets

-8.9      Delete analysis and design funds already spent by DOE  
-----

\$ 84.3M      Future LSS development and operation budget through FY 2000

## LSS Concept/Design Alternatives Summary

	Concept/Design Alternative	LSS Budget Reduction* (\$M)		LSS Benefit Reduction	Impact on the LSS Rule			Recommended by Working Group
					None	Minor	Major	
1.	Capture DOE LSS material using INFOSTREAMS versus LSS	73.6		None		■		Yes
2.	Adapt INFOSTREAMS search and retrieval capability to LSS	<1.0		None	■			Yes
3.	Remove LSS telephone connect charges from the LSS Budget	1.8		None	■			Yes
4.	Remove local LSS costs (workstations/other) from the LSS Budget	3.1		None	■**			Yes
5.	Remove costs to "pre-process" material for input to the LSS	3.0		None	■			Yes
6.	Eliminate one of two training/user support locations	2.5		Small	■			No
7.	Use existing infrastructures (rather than LSS) provide paper copies of LSS documents on demand	3.9		Small		■		No
8.	Minimize non-DOE capture costs:							
	a. DOE capture all non-DOE LSS material; or	14.0		None			■	See text
	b. LSS capture non-DOE LSS material using capabilities adapted from INFOSTREAMS		<1.0	None	■			Yes***
9.	No on-line images/limited on-line images							
	a. Digital images of all pages stored off line in LSS; paper copies of requested pages mailed to individual users; or	-1.3		Large			■	No
	b. No digital images (off line or on line) in LSS; microform images of all pages sent to key user locations; or		3.1	Large			■	No
	c. Digital images of "graphic" pages only stored on line in LSS LSS; digital images of text pages sent to key user locations		-0.8	Medium			■	No
10.	Eliminate some documents from LSS database and place them in NRC/DOE public document rooms in a timely manner:							
	a. Establish a cutoff date unless to be relied upon; or	0.3 per million pages excluded		Small			■	No
	b. Include only "core documents" in the database			Large			■	No
11.	Reduce early availability of LSS database:							
	a. Limit use and size of database from early 1996 until early 1998; or	3.3		Small	■**			Yes
	b. Don't make database available for use until early 1998		5.7	Medium	■**			No

\*Combination of cost savings and budget reallocations, FY 1991 - 2000

\*\*Modify expectations of LSS participants

\*\*\*Not required if 8a is adopted

A detailed discussion and evaluation of each of the proposed alternatives shown in the table is presented below:

1. Capture DOE LSS material using INFOSTREAMS versus LSS

DOE has proposed to use the capabilities of its INFOSTREAMS system to convert its LSS documents into electronic format that will be compatible with the needs of the LSS. This processing would be in lieu of similar processing that was originally planned as a component of the LSS and included in the SAIC cost estimate.

The Working Group supports the concept of substituting INFOSTREAMS processing of DOE's LSS material for LSS processing since DOE has agreed to meet both LSS standards and production schedules. Using INFOSTREAMS, DOE will unitize materials, prepare a full header for each unit, eliminate duplicates, set record pointers, maintain audit trails, clean up ASCII text, create bit-mapped images and meet any other LSS capture requirements. No indexing or other capture functions (beyond quality assurance, document correction and hearing support) will need to be performed in the LSS for DOE's LSS material. Also, Infostreams annual production volumes for LSS material will be reasonably consistent with the expected quality assurance and database loading capacities/requirements of the LSS, so that over the 1995-2001 timeframe, material can be quality assured and the database can be loaded six months before the receipt of DOE's license application. If prioritized loading of specific materials is deemed to be a requirement, INFOSTREAMS will produce documents consistent with such a loading schedule. Only a minor change would be required to the LSS rule and there would be no loss of benefit to LSS users if this alternative is adopted.

Using INFOSTREAMS to process DOE's LSS materials will result in a significant shift of costs from the LSS budget to OCRWM's internal IRM budget. The resultant reallocation from the LSS budget if this alternative is adopted would be \$73.6 million. This is the estimated cost for processing these documents in the LSS (\$76.3M) less the costs that must be incurred by the Office of the LSS Administrator in FY 1994-FY 1995 to receive this electronic data from DOE and assure its quality prior to its entry into the LSS database (\$2.7M).

RECOMMENDATION: ADOPT

2. Adapt INFOSTREAMS search and retrieval capability to the LSS

Although the INFOSTREAMS search and retrieval software (both commercial off-the-shelf and application software) has not yet been specified by DOE, the capabilities being examined by DOE appear to be powerful and easy to use. Since there is reason to believe that this software could be used in the

LSS, albeit with some modification, this alternative has the potential for reducing LSS software development costs. DOE plans to develop upgraded search and retrieval software (for both text and image) on a scale and schedule that will permit its timely evaluation for use in the LSS. The estimated cost savings for this alternative are difficult to assess because the extent of the required modifications to the INFOTREAMS software are not known, but the savings would be \$1.0 million or less.

RECOMMENDATION: ADOPT

3. Remove LSS telephone connect charges from the LSS Budget

Although LSS telephone connect costs (\$1.8 million) were included in SAIC's LSS program costs, the LSS rule makes telephone connect charges the responsibility of individual LSS participants. Therefore, they should not be included in future LSS budget estimates.

RECOMMENDATION: ADOPT

4. Remove local LSS costs (workstations/other) from the LSS Budget

The LSS rule makes individual participants responsible for the cost of their "individual computer facilities to have remote access to the LSS (10 CFR 2 §2.1007)." This includes the costs of all hardware and software needed to connect to the LSS wide area network and to search the LSS database. LSS workstations were included in SAIC's LSS program costs, at a cost of \$3.1 million for their purchase and maintenance.

Given that the LSS rule makes LSS participants responsible for the cost of these workstations, their cost should not be included in future LSS budget estimates. Even though participants would pay for their workstations, it may be advantageous for them to be purchased through the LSS contract.

DOE and NRC will ensure that adequate terminal access facilities are provided at the public document rooms.

RECOMMENDATION: ADOPT

5. Remove costs to "pre-process" material for input to the LSS

The LSS rule makes all LSS participants responsible for "pre-processing" their material for submission to the LSS where it would be processed and then loaded into the LSS database. The \$193 million baseline includes \$8.1 million for LSS participants' "pre-processing." Of the \$8.1 million, \$5.1 million was removed from the LSS budget by virtue of it being included in the \$73.6 million removed

from the LSS budget by Alternative 1. The other \$3.0 million is removed under this alternative.

RECOMMENDATION: ADOPT

6. Eliminate one of two training/user support locations

The LSS Administrator's current plan for the LSS is to have both Eastern U.S. and Western U.S. training facilities and user support functions to train and assist end users. By doing so, these functions will be provided at locations and times convenient for most users, regardless of their geographic location. Under this alternative, the Eastern U.S. training and support operation would be eliminated, which would save \$2.5 million. The Western U.S. operation would be retained at its currently projected size, by concentrating more on training the trainers, than on training end users. LSS participants in the Eastern U. S. would have to travel greater distances to be trained, perform more of their own end-user training and would receive help desk support only between the hours of 10:30 a.m. and 6:00 p.m., Eastern time.

The Working Group does not recommend adopting this alternative. High quality training and user assistance are both viewed as vitally important functions to the success of the LSS. Eliminating the Eastern U. S. training and support operation would result in a significant reduction in services. The savings associated with this alternative are not sufficient to warrant the reduction in services to LSS participants and to the public. Moreover, projected savings could be offset substantially by the additional burdens placed on the system by insufficiently trained and supported users.

RECOMMENDATION: DO NOT ADOPT

7. Use existing infrastructures (rather than LSS) to provide paper copies of LSS documents on demand

While paper copy distribution of agency documents are inherent responsibilities of DOE and NRC (through the Freedom of Information Act and PDR activities), the LSS rule (10 CFR 2 §2.1007) specifies that LSS users can order paper copies on-line and expect to receive them without charge from the LSSA. This alternative considers shifting the responsibility back to the organizations that produced or acquired the information stored in the LSS.

The Working Group does not recommend adopting this alternative. Although it might seem that the use of existing infrastructures would be more cost effective than creating a new function within the LSS, the cost to DOE, NRC and others would likely be more than the \$3.9 million projected for the LSS. DOE, NRC and others would each have

to provide the equipment, staff and other resources required to generate paper copies from the LSS database. Thus performing this function on a decentralized basis (and on a smaller scale at each site) would likely increase costs rather than reduce them.

RECOMMENDATION: DO NOT ADOPT

8a. Minimize non-DOE capture costs - DOE capture all non-DOE LSS material

Although DOE has proposed to electronically process its LSS material through INFOSTREAMS (about 80 - 86 percent of all LSS material), there is another 14 - 20 percent from NRC and other LSS participants that must also be electronically processed. The SAIC design called for this processing to be done in a central capture facility operated by the LSS Administrator. This alternative considers having DOE process NRC's and other participants' LSS material through INFOSTREAMS.

The Working Group sees this as a technically feasible way to reduce costs. Given that INFOSTREAMS can mirror LSS processing requirements and will be doing so on a very large scale, it would be more economical for DOE to assume responsibility for the other 14 - 20 percent than to establish a separate small scale operation in the LSS to do so. If this alternative were adopted, the LSS could be limited to document receipt, quality assurance, hearing support, database loading and information dissemination and the LSS budget could be reduced by \$14 million. This alternative would remove from the LSS budget the \$9 million cost of labor, facilities, equipment and maintenance for processing NRC and other participants' materials. In addition economies of scale to be achieved in INFOSTREAMS may permit DOE to perform this processing for less than the cost of doing it in the LSS. The \$5 million cost of designing, developing, testing, operating and maintaining LSS capture software and procedures could also be eliminated, except for a small amount needed for quick turn around information capture during hearings. DOE would incur the cost of processing non-DOE material and the costs of any changes to INFOSTREAMS software and procedures necessary for processing it. The Working Group believes that adequate quality checks can be instituted by both the LSS Administrator and LSS participants to assure accurate processing of non-DOE material through INFOSTREAMS.

With regard to this alternative, DOE has serious concerns about any plan that would give DOE responsibility for entering other parties' submissions. DOE's access to, and control over, other parties' materials was very contentious during the negotiated rulemaking. Furthermore, there are serious policy and operational questions related to responsibility for intake prioritization, liability for



accuracy and timeliness of entry of other parties' materials, and budgeting and procurement of incremental resources. DOE's level of concern is such that it can not concur in recommending this for any further consideration.

Implementation of this alternative would require the agreement of LSS participants and a change to the LSS rule.

RECOMMENDATION: ADOPT PROVIDING DOE AND OTHER PARTICIPANTS CONCUR

8b. Minimize non-DOE capture costs - LSS capture non-DOE material using software adapted from INFOSTREAMS

This alternative would retain the central LSS facility to process documents submitted by non-DOE participants. Software development costs could be reduced somewhat by adapting INFOSTREAMS developed software for LSS use.

The Working Group finds this to be an acceptable alternative to the current plan, but lacking the distinct advantages of 8a above. The only savings that could be realized from this alternative would be the amount of development cost avoided by adapting INFOSTREAMS developed software to the LSS (estimated at \$1 million or less).

RECOMMENDATION: ADOPT IF ALTERNATIVE 8a IS NOT ADOPTED

9. No on-line images/limited on-line images

The LSS rule requires that digital images of all pages be stored on-line in the LSS. In the SAIC design, the LSS would have duplicate image servers in both the Eastern and Western part of the U.S. to minimize telecommunications costs and to have one set of optical disks serve as backup to the other. The benefits of on-line images are three-fold; users will be able to view images of document pages and technical data on screen with little delay; capture system staff will access page images to verify that duplicate documents are not being entered into the database and image servers will be linked to high-speed printers to produce hard copy of document pages upon request.

LSS users will search the LSS text and bibliographic header databases in order to locate relevant documents; however, they will need to view images in order to see document pages in their original format, including italics, underlining, highlighting and strikeouts. More importantly, images of document pages provide access to "graphic" information elements that cannot be captured in text and bibliographic header databases. Such elements include graphics, line drawings, signatures, mathematical equations, chemical formulas and marginalia.

Many of the technical reports to be included in the LSS will

contain large numbers of charts and graphs (often as much as 50 percent of the pages) that the user must see in order to adequately review and understand the reports. Since the "graphic" information is typically interspersed with text, the ability to view both text and graphical information without interrupting ongoing work to go to other sources will be very beneficial for users. By one estimate, half of the LSS material will eventually consist of pages containing graphic information. Some technical users have asserted that access to on-line digital images is an important attribute of the LSS and that eliminating this capability would impair its use in the technical review process.

Having on-line images available is especially important in the case of "technical data" because very little of this material has text that can be captured and stored in the LSS. Technical data is largely raw data compiled during scientific investigations. This data consists of items such as handwritten notes, graphs, maps, photographs, sketches, numeric tables and computer tapes. The LSS will contain bibliographic indexes for this material and images for all of it that can be scanned. The LSS will contain a large volume of this material; the Yucca Mountain Project Office already has a backlog of roughly four million pages of this material.

A number of cost and benefit issues should be examined carefully when considering limiting or eliminating access to on-line images from the LSS. There are several cost components associated with providing on-line images in the LSS -- the cost of creating digital images from source documents, the cost of storing and maintaining those images in the LSS database, the cost of workstations, and the cost of communications services to access the images. The cost of creating the images from source documents is not an issue for DOE's LSS material (80 - 86 percent of total LSS material), because DOE plans to create digital images of its LSS material for its own internal purposes using INFOTREAMS irrespective of LSS requirements. The cost to create these images is included in the previously discussed \$73.6 million estimate for processing DOE's LSS material through INFOTREAMS (see Alternative 1.) Storing and maintaining images in the LSS database is relatively inexpensive, because of the high storage density of optical disks. Image storage and maintenance costs for the planning period total \$2.4 million for 20 million pages of LSS material. On-line access to images requires properly configured workstations and increased communications traffic. These are both user costs because the LSS rule requires users to pay for their own workstations and telephone connect charges (see Alternatives 3 and 4). Therefore any savings in workstation and communications costs due to limiting or eliminating on-line images from the LSS accrue to users, not the LSS.

Eliminating or reducing access to the LSS negates or reduces many of the benefits discussed above. Because of the large size of the LSS database, it would be awkward and time consuming for users to view images of documents and technical data in an off-line environment. Manually finding particular images or groups of images in a collection of tens of millions of pages of paper, or tens of thousands of microfiche, or thousands of CDROM disks is not practical, particularly considering how un-useful an information collection becomes when people are continually withdrawing, returning, misfiling (or never returning) items of interest. Because of this difficulty, those who can, will likely develop their own LSS image access systems if the LSS eliminates or reduces access to online images. This would result in a large duplication of effort and in most cases, additional costs to the Nuclear Waste Fund.

9a. No on-line images - Digital images of all pages stored off line in LSS; paper copies of requested pages mailed to individual users

If this alternative is adopted, no on-line access to images through the LSS would be available. Instead, paper copies of the images on these disks would be generated and mailed to LSS participants upon request.

The Working Group does not recommend adoption of this alternative for two reasons. First, adopting this alternative would increase the cost of the LSS by \$1.3 million according to the SAIC benefit-cost analysis. The savings potentially achieved by eliminating equipment and software for on-line image storage are more than offset by the increased cost of providing voluminous paper copy production from optical disks to LSS participants. Perhaps more important, however, is the large negative affect it may have on LSS participants. They would have to deal with both ordering and maintaining huge quantities of paper and would experience difficulty locating information in a paper environment.

RECOMMENDATION: DO NOT ADOPT

9b. No on-line images - No digital images (off line or on line) in LSS; microform images of all pages sent to key user locations

This alternative would replace on-line digital images with off-line microform stores located at key user locations. Having located documents of interest using the text search capabilities of the LSS, users would visit user microform files at key locations to locate and view document images.

The Working Group does not recommend adopting this alternative. In their benefit-cost analysis, SAIC concluded

that \$3.1 million could be saved by this option, but that there would also be a large reduction in benefit to LSS users. The LSS is currently expected to hold several million pages of images (maps, technical charts, etc.) that are either non-textual or only partly textual. The only way users could view these images would be away from their LSS workstations, perhaps in other buildings. This could be disruptive and would remove one of the most beneficial functions of the LSS design.

RECOMMENDATION: DO NOT ADOPT

9c. Limited on-line images - Digital images of "graphic" pages only stored on line in LSS; digital images of text pages sent to key user locations

Under this alternative images of pages that consist of text with no marginalia or embedded graphics would not be stored on line in the LSS, but images of all other LSS material would. Optical disks containing digital images of "text" pages (no marginalia or embedded graphics) would be sent to key user locations. Based on information from the SAIC communications design document (Ref. 9), it appears that there are about 40 such sites. Users would want to be able to view images of "text" pages because of the occurrence of errors. Information in the text and header databases will always be less than 100 percent due to errors introduced during the conversion and cataloging processes. Some users will also want to view these images in order to see text in its original format. The latter is important because affordable ASCII text conversion does not preserve features such as type face and size, italics, underlining, highlighting and strikeouts.

If images of "text" pages were loaded onto 600MB CDROM disks for distribution, about five new CDROMs could be sent to each key location each week, assuming that LSS information is made available at a rate of 4 million pages per year and 50% of it consists of text with no marginalia or embedded graphics. Thus each key user location would amass a collection of 1,250 CDROMs by the time DOE submits the license application in 2001. The cost to produce the CDROMs and distribute them to the 40 sites would be about \$1.4 million. A set of CDROMs and a CDROM reader would be required for many standalone image workstations. At sites with multiple workstations, the CDROMs could be stored in jukeboxes connected to networks to permit access by multiple users.

It would at first appear that this alternative would cut LSS image storage and maintenance costs in half and reduce LSS telecommunications costs. The amount of reduction in telecommunications costs would be equal to the telephone connect charges for transmitting images of pure text pages if they were available on line. However, LSS participants

must pay for their telephone connect charges, so limiting on-line images will not reduce future LSS telecommunications costs.

Storage and maintenance cost savings are smaller than expected because a full set of optical images must be kept to support LSS functions. In order to provide paper copies on demand, as described in the LSS rule, in an efficient manner and in order to assure accurate duplicate data entry checking, pages of all documents must be maintained on line on optical disks to support these functions at the Western U.S. site. Half (\$0.6 million) of image storage and maintenance costs for the Eastern U.S. image storage facility could, however, be eliminated from the LSS budget.

The Working Group does not recommend adoption of this alternative. It represents a negative savings (increase) to the LSS budget of \$0.8 million. It offers no additional communications cost savings and image storage and maintenance cost savings (\$0.6 million) would be more than offset by the \$1.4 million cost of producing and distributing CDROMs.

In addition, participants (and ultimately the Nuclear Waste Fund in most cases) would bear the cost of CDROM readers, jukeboxes and any networks necessary for accessing the CDROMs. A CDROM reader for a single workstation could be purchased and maintained over the planning period for as little as \$1,500; however, since images would be loaded onto CDROMs in the order that documents are added to the LSS, a workstation user would likely have to load multiple CDROMs in order to view the images of the several documents found in a particular search or series of searches. This could require multiple trips (across the room, down the hall or wherever) to retrieve the appropriate CDROMs from the collection. This procedure would be further prolonged in situations where the CDROMs were already in use elsewhere. To avoid this problem users could share their set of LSS CDROMs among a number of workstations by purchasing and maintaining appropriate jukeboxes at an estimated cost of \$0.6 million per site. Doing this at multiple sites would, of course, duplicate effort and cost.

**RECOMMENDATION: DO NOT ADOPT**

- 10a. Eliminate some documents from LSS database and place them in NRC/DOE public document rooms in a timely manner - Establish a cutoff date unless to be relied upon

A fundamental assumption in the LSS concept is to have in the electronic database all material that LSS participants agreed might be relevant or potentially relevant to the licensing of the repository, irrespective of date. This alternative assumes that a cut-off date would be established, such that low-value (older) material would not

be entered into the LSS but would be available to participants in paper or microform, unless a party intended to rely on one or more of these documents. The Working Group recommends that this alternative not be adopted because there is little potential for achieving additional cost savings. Most of the potential savings that could be achieved are in the area of reduced capture costs, which will not be included in the LSS if Alternatives 1 and 8a are adopted. The only additional costs that could be eliminated would be a portion of on-line document storage cost. Because that cost is only \$0.3 million per million pages, agreement would have to be reached to eliminate a large number of pages from the database in order to achieve an appreciable savings. The idea of establishing a cut-off date was discussed during negotiations on the LSS rule but an agreement was never reached.

RECOMMENDATION: DO NOT ADOPT

- 10b. Eliminate some documents from the LSS and place them in NRC/DOE public document rooms in a timely manner - include only "core" documents in the database

This alternative would limit the size of the LSS database by defining a core set of documents for inclusion in the database. Indexes of all LSS material would be made available from time to time to help parties determine what additional documents they might want to have added to the core database.

The concept of limiting the scope of the LSS database was discussed during the LSS negotiated rulemaking; however, it was concluded that the LSS needed to have a comprehensive database of searchable full text, constrained only by the Topical Guidelines and certain exclusions described in the LSS rule. The argument was made that a complete and comprehensive electronic database gives users a very high confidence level that all relevant or potentially relevant documents can be readily located and examined.

The Working Group recommends that this alternative not be adopted. Similar to Alternative 10a, at \$0.3 million per million pages, a large number of pages would have to be eliminated from the database in order to achieve appreciable additional savings. Moreover, the process of agreeing on which documents to eliminate from the database would be so subject to interpretation and debate it is questionable whether a large number of pages could actually be eliminated.

RECOMMENDATION: DO NOT ADOPT

- 11a. Reduce early availability of the LSS database - limit use and size of database from early 1996 until early 1998

In the LSS Administrator's \$193 million estimate, he planned to give 30 users access to a small number of critical documents (approximately two million pages) in early 1996 and then expand the number of users and the database starting six months later. Alternative 11a would provide limited access to a database of approximately two million pages for 18 months beginning in late 1996<sup>1</sup>; full access would not begin until early 1998, three and one half years before DOE is expected to submit its license application to the NRC.

The Working Group recommends adoption of this alternative because of the operational cost savings and because all participants would be able to have full access to the database starting three and one-half years before the license application submittal date, which was the expectation at the time the LSS rule was promulgated. Although there is a small reduction in benefit, the magnitude of the cost savings (\$3.3 million) make it attractive.

RECOMMENDATION: ADOPT

11b. Reduce early availability of the LSS database - don't make database available for use until early 1998

This alternative makes further operational savings possible since no database would be available to anyone until three and one-half years before DOE is expected to submit its license application. This alternative would provide no user access until early 1998, although the system would have to be available to start loading a year to 18 months beforehand.

The Working Group does not recommend the adoption of this alternative. The increase in savings over Alternative 11a is significant (\$3.1 million), but so is the benefit lost by not having high priority material accessible during the 1996/1997 timeframe. Moreover, the experience gained from use of the system on a small scale during 1996/1997 could be invaluable to the future success of the LSS under full load conditions.

RECOMMENDATION: DO NOT ADOPT

Other Considerations

In addition to the measures recommended by the Working Group, there are other future considerations that could lead to a reduction in the cost to develop and maintain the LSS or a reduction in the cost to

---

<sup>1</sup> This is the earliest possible data assuming a decision is made to accept the Working Group's recommendations by the end of March 1992.

capture documents through INFOSTREAMS. Advances in decision based software systems and improved text searching algorithms also show promise for reducing the processing costs associated with the LSS. The cost impact of new and more sophisticated search and retrieval software products that could potentially reduce the need (at a very substantial cost reduction) for meeting 99.8% accuracy of ASCII text when processing LSS documents for full text are significant. Such software would eliminate much of the high clerical costs to "condition" documents to raise their quality to a level acceptable for present search algorithms to be effective.

The earlier section reflects a series of alternatives for reallocating or reducing costs based on reuse of DOE's INFOSTREAMS technology, reduced functionality, reduced availability, and other strategies. Many of the items from the chart entitled LSS Concept/Design Alternatives Summary are interdependent or could be considered in conjunction with others as part of a "package". In this section, several discrete strategies are presented.

Sample Strategy #1: Is it feasible that not all documents be included with searchable full text, but rather made available via bibliographic header and bit-mapped images only?

Text conversion is the single most costly element of all LSS processes. It was incorporated into the LSS design as a blanket requirement for all documents: 1) before the header fields were decided, 2) as a response to legal representatives who were familiar with the technology, and 3) recognizing that subject cataloging had inherent deficiencies.

Conversely, the text of some documents adds little or nothing to their retrievability if they have been competently and fully cataloged. A bibliographic header does provide search and retrieval capabilities and is an appropriate level of treatment in some circumstances. By using bibliographic headers, where the associated image is available on line, participants still have access to those materials.

Some situations typical of the LSS document collection are amenable to differential treatment:

Transmittal memos and letters attached to reports, studies, etc., are often not content rich. Rather, it is the item being transmitted that contains the information of value. So, if attachment relationships and cross-reference fields are properly designed, and if the item attached, itself, is full-text searchable the entire package (transmittal and report) is still eminently retrievable via text search.

Another situation consists of the flip charts and other presentation materials which are attached to textual meeting minutes.

DOE's contribution to the LSS holdings is estimated to be 80% of the low-volume estimate and 86% of the high-volume estimate. Of DOE's contribution for the low volume estimate, 4,320,000 pages (65% of all



its documents but only 12% of its pages) will be correspondence (letters, memoranda, telex, etc.). A simple analysis of impact for not including text for all such correspondence follows:

Reduce all OCR intake labor for this proportion of the material:

OCR Pre-processing (\$10,136,000 x .12)	\$ <1,216,320>
OCR Cleanup (\$16,650,000 x .12)	<1,998,000>
Scan/Text Supervisors (\$4,864,000 x .12)	< 583,680>
Disk Storage (\$5,120,000 x .12)	< 614,400>

Increment for offsetting increase in hardcopy printout from image (\$9,045,000 x .12)	<u>1,085,400</u>
--	------------------

Net True Savings for Bibliographic Header & Image,  
but no full text for DOE correspondence:      \$<3,327,000>

This is a simplistic presentation insofar as storage would not really decrement proportionately, since text from correspondence is less character-dense than equivalent pages of reports and publications. And, it is more palatable if text is omitted only for correspondence that is attached to a text-searchable report, resulting in a smaller percentage reduction. However, it is representative of strategies focused on the peculiarities of the document collection and knowledge of users' retrieval expectations.

Sample Strategy #2: What economies could result if data accuracy requirements are reduced to 98% accuracy rather than 99.8% because "intelligent" retrieval software compensates?

DOE's LSS Prototype showed that the most accurate OCR device tested achieved an average character accuracy of 98.6%, which corresponds to 25 errors on an average 1800 character LSS page.<sup>2</sup> In that same prototype, it was found that text accuracy must approach 99.8% or users would lose confidence that they were able to retrieve all critical documents, and thus lose confidence in the LSS itself. Under SAIC's design, documents for which the OCR output accuracy was not in the 95-98% range would require additional editing that would exceed the cost of a complete, manual rekey of the entire page.<sup>3</sup> It was also found that, on average, editing represents from 65-75% of the total cost of text conversion, and that correcting OCR-induced errors constitutes 67% of that total editing cost. The multiple OCR device approach for intake, reflected in SAIC's final cost estimates, was based on analyses showing that 2/3 of OCR-induced errors could be eliminated by merging and matching streams from multiple OCR devices.

DOE is studying content/concept based retrieval software to augment classic Boolean tools. The most significant aspect of this software is that it profiles a document's content. But, one unanswered

---

<sup>2</sup> Dickey, Lois. "Operational Factors in the Creation of Large Full-Text Databases", INFOTECH '91. p.41.

<sup>3</sup> Ibid., p. 45.

question is how much of the document must be "read" before all the relevant terms and topics have been identified?

If, after analyzing the first 20 pages (@99.8%) of a 350 page report, the software "knows" what the document is about, isn't the rest of the document superfluous in adding to our understanding of its content? If after these 20 pages our matrix is already "saturated", we have 330 pages where the input accuracy of the text could be as low as 90% and have no impact whatsoever on our ability to characterize the document.

What if the entire document were 90% accuracy? -- then it may take an additional 10 pages of text to find a "clean" occurrence of terms and topics before the matrix was again saturated. But, again, 320 succeeding pages contribute nothing more to our understanding.

In a way, the "intelligence" of the software compensates for typographical errors by having access to enough bulk ASCII, with enough clean text, to be able eventually to correctly characterize the document. DOE's testing still has to validate the concept. For example, DOE does not know where the "saturation" level is, and if it is affected by the overall length of a document. How much "clean" ASCII is needed? Will it work as well on an eight page letter as it will on the longer report? Will a high percentage of uncorrected ASCII result in an unacceptable level of false characterizations, resulting in associated false drops during retrieval? How would the highlighting of occurrences of terms in text be implemented in a "dirty ASCII" environment? Can the matrix compiled during the filtering of incoming text be added somehow to a simple bibliographic header with associated image, obviating the need for text?

Academic papers about the new software packages which utilize "fuzzy logic" indicate that this approach will work for search and retrieval -- and this would be sufficient because it is the images and not ASCII that are relied on for introduction as exhibits. Will LSS users be satisfied that such intelligent software is able to compensate for typographical errors? If they could be convinced, and, if one of these new software packages is roughly comparable in cost to a current state-of-technology, Boolean-based package such as BASIS+, then the following scenario could apply: we could remove the multiple OCR devices from SAIC's final design, and accept the basic 98% text accuracy with no additional text editing and OCR cleanup.

A simple calculation of savings is as follows:

5 Capture Systems requiring less OCR hardware:	\$< 500,000>
Eliminate OCR compare software:	< 100,000>
Eliminate OCR Cleanup Staff:	<16,650,000>
Eliminate OCR Cleanup Supervisors:	< 1,200,000>
1 Correction Station's Text Cleanup Eliminated:	< <u>1,050,000</u> >
Total Savings for 98% accuracy:	\$< <u>19,500,000</u> >

A demonstration of this strategy was made to DOE utilizing commercially available off-the-shelf technology, EXCALIBUR software, which is VAX compatible but does not work in conjunction with DOE's current BASIS+ records management software.

Both of the approaches outlined above would represent major deviations from what the parties agreed to during the Negotiated Rulemaking, and they would all be contentious to varying degrees.

Other strategies are conceivable: if bibliographic headers could be made to store the entire subject-content matrix of every textual document (derived from DOE's expert system software), the OCR and text analysis processes would not decrease, but perhaps no text at all would have to be stored or retrieved. This would have major impacts on the amount of disk storage, the size of the search engine hardware, the organization of databases (no partitioning), database loading and maintenance, the size of the telecommunications lines, etc. The cost ramifications of such strategies would require robust, detailed, and professional feasibility and benefit-cost studies outside the scope of this paper.

#### Observations vis-a-vis Future Actions on the LSS

The LSS Administrator identified two major issues confronting the LSS program (Ref 14). These are the budget responsibility (DOE or NRC) and the schedule for the development of the LSS. The issue of which organization (DOE or NRC) builds the LSS raises a number of policy questions. The Working Group notes that if a decision is made to assign this responsibility to NRC, there may be opportunities for cost optimization by joining the development of the LSS and NRC's internal document management systems.

The NRC's Nuclear Document Control System (NUDOCS) serves as NRC's document storage and retrieval system and provides bibliographic reference material to the PDR. Many of the functions required by the LSS and planned for inclusion in INFOTREAMS are also useful to the NUDOCs system. A combined INFOTREAMS/LSS/NUDOCS information and technology exchange (if the time schedules are complimentary) could provide even further cost reductions. The amount of these cost savings cannot be estimated at this time and would be achieved in other DOE or NRC budget categories (not LSS). However, such an exchange would undoubtedly offer cost savings since many of the requirements of all three systems are quite similar. The Working Group supports further efforts to explore the development of a combined DOE/NRC working relationship that fosters the exchange of technology related to automated document processing.

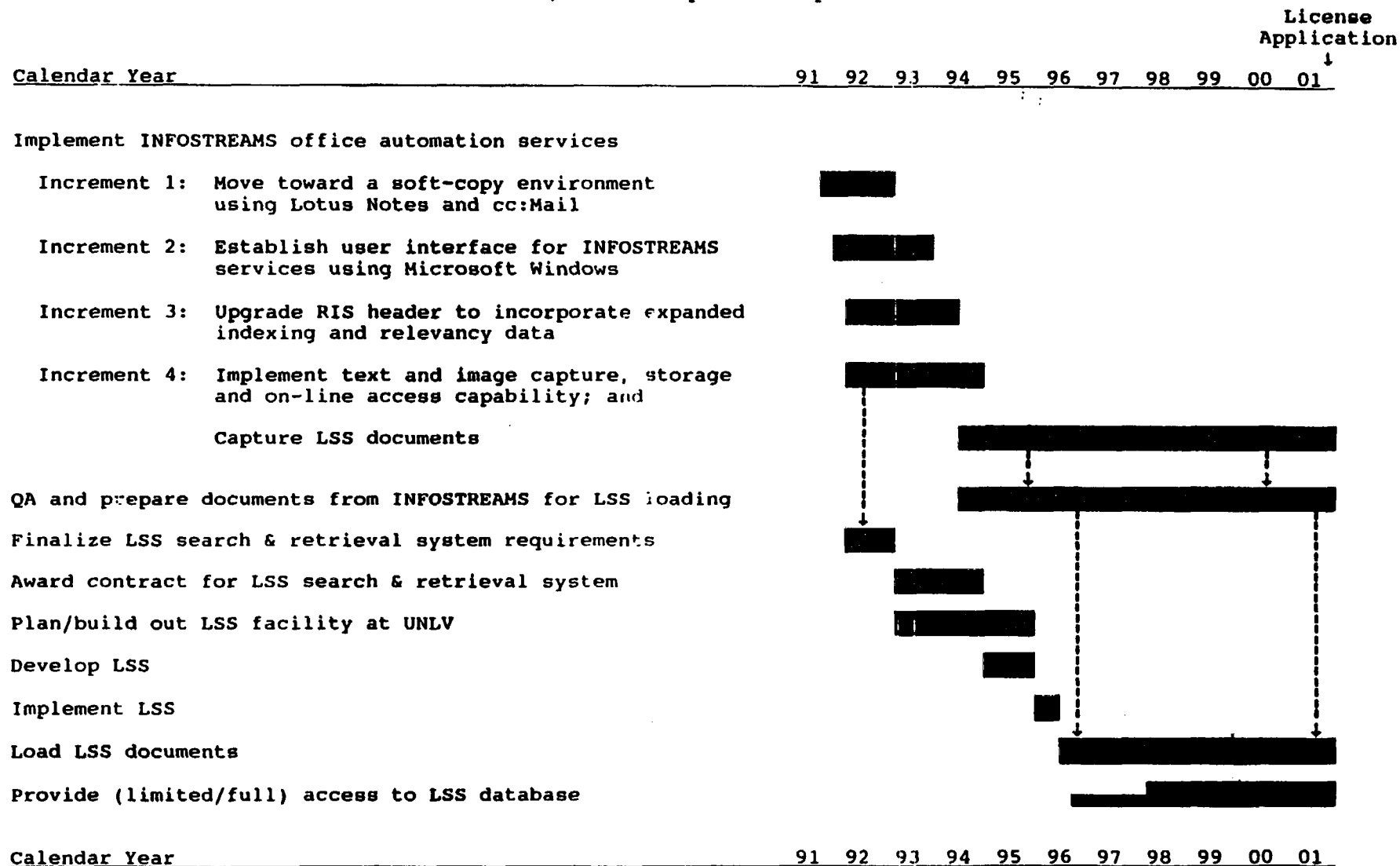
The Working Group also notes that if repository development progresses on the schedule provided by DOE (Refs. 30-31) and if an LSS of the SAIC design is to be used, LSS development activities should resume now if the system is to be available to support the NRC staff's technical review and pre-hearing discovery, albeit on a small scale starting in 1996/1997.

It normally requires about five years to procure, develop, implement

and test a major automated system like the LSS. If the LSS is limited to essentially a search and retrieval capability (the data being captured using INFOSTREAMS), the procurement will be simpler and perhaps the five-year timeframe can be reduced to four years. Following system testing, an initial quantity of high priority documents must be loaded prior to providing access to the system. Database loading would continue for about four years following the initial "high priority" loading and users would gain access to increasing amounts of material, with the database expanding to about 20 million pages by six months prior to the submission of DOE's license application in 2001. If the system is not developed on a schedule that makes it available for database loading four to five years before the submission of the license application, it is unlikely that the estimated 20 million pages of relevant material would be in the LSS system by 2001.

Attachment 1 provides a schedule for development of the LSS assuming the use of INFOSTREAMS as the data capture mechanism and the LSS as a search and retrieval system. This schedule assumes a decision is made to proceed with the LSS by the end of March 1992.

# INFOSTREAMS/LSS Development & Implementation Schedule\*



\*Assumes a decision is made by the end of March 1992 to accept the Working Group's recommendations

## REFERENCES

Note: A number of these references are pre-decisional in nature and are not publicly-available documents.

1. J.W. Bartlett (DOE) letter to I. Selin (NRC) dated August 23, 1991.
2. I. Selin (NRC) letter to J.W. Bartlett (DOE) dated September 11, 1991.
3. G.F. Cranford memorandum for H.L. Thompson on "Schedule for Technical Review of Licensing Support System" dated September 20, 1991.
4. Licensing Support System Preliminary Data Scope Analysis, U.S. Department of Energy Report DOE/RW-210, January 1989.
5. Licensing Support System Preliminary Needs Analysis, U.S. Department of Energy Report DOE/RW-212, January 1989.
6. Licensing Support System-Level Requirements Document, Science Applications International Corporation Report dated December 12, 1990.
7. Licensing Support System Capture System Design Document, Science Applications International Corporation Report dated March 9, 1989.
8. Search and Image Design Document for the Licensing Support System, Science Applications International Corporation Report dated November 16, 1990.
9. Communication System Design Document for the Licensing Support System, Science Applications International Corporation Report dated November 16, 1990.
10. Licensing Support System Benefit-Cost Analysis, Science Applications International Corporation Report dated January 1989.
11. Revised Cost Analysis for the Licensing Support System, Science Applications International Corporation Report dated December 10, 1990.
12. L.J. Donnelly (LSSA) memorandum for the Commission on "Licensing Support System Program and Budget Responsibility" dated June 19, 1991.
13. K.M. Carr (NRC Chairman) memorandum for Commissioners Rogers, Curtiss and Remick on "Licensing Support System Program and Budget Support Responsibility" dated June 24, 1991 (with associated comments).

14. L.J. Donnelly (LSSA) note for Chairman Selin dated July 10, 1991.
15. L.J. Donnelly (LSSA) memorandum for The Chairman and Commissioner Curtiss on "Follow-up to August 20, 1991 Meeting on the Licensing Support System" dated August 22, 1991.
16. L.J. Donnelly (LSSA) memorandum for LSS ISC Members on "LSS Internal Steering Committee Review of LSS Needs and Functionality" dated March 18, 1991.
17. R.M. Bernero (NMSS) memorandum for L. Donnelly (LSSA) on "Consideration of Distributed Databases in Cost Consideration for the Licensing Support System" dated March 18, 1991.
18. L.J. Donnelly (LSSA) memorandum for R.M. Bernero (NMSS) on "LSS Internal Steering Committee Review of LSS Needs and Functionality" dated March 22, 1991.
19. "LSSA's Proposed LSS Procurement Strategy" presented to the LSS Advisory Review Panel on July 17, 1991.
20. "LSSA's Proposed LSS Development Schedule" presented to the LSS Advisory Review Panel on July 17, 1991.
21. "Submission and Management of Records and Documents Related to the Licensing of a Geologic Repository for the Disposal of High-Level Radioactive Waste," Federal Register, Vol. 54, No. 71, pp. 14925-14955 (April 14, 1989).
22. "Procedures Applicable to Proceedings for the Issuance of Licenses for the Receipt of High-Level Radioactive Waste at a Geologic Repository," Federal Register, Vol. 56, No. 38, pp. 7787-7799 (February 16, 1991).
23. "SAIC Examples of LSS Design Efforts" provided to the LSS Working Group by the DOE on October 23, 1991.
24. "Leveraging Technological Developments -- Ongoing Evaluations by DOE" presented to the LSS Working Group by the DOE on October 24, 1991.
25. "Status of Office of Civilian Radioactive Waste Management Information Management Environment and Licensing Support System" Presentation by Barbara Cerny (OCRWM) to the Commission dated August 20, 1991.
26. "Levels and Views: An Architecture Definition Methodology" Presentation by Donna Sitterson (TRW) to LSS Working Group dated September 26, 1991.
27. L.J. Donnelly (LSSA) memorandum for G.F. Cranford on

"Proposal for Evaluation by LSS Technical Working Group" dated October 3, 1991.

28. "A Low Cost Alternative to the LSS" presented to the LSS Working Group by the LSS Administrator on September 17, 1991.
29. B.J. Youngblood (NMSS) memorandum for G. Cranford on "LSS Technical Working Group" dated October 15, 1991.
30. Radioactive Waste Management System Project Decision Schedule, Revision 1, U.S. Department of Energy Office of Civilian Radioactive Waste Management Report DOE/RW-0310P, June 1991.
31. Draft Mission Plan Amendment, U.S. Department of Energy Office of Civilian Radioactive Waste Management Report DOE/RW-0316P, September 1991.





**Department of Energy**  
Washington, DC 20585

January 10, 1992

Mr. Gerald Cranford  
Director, Information Resources Management  
U.S. Nuclear Regulatory Commission  
Washington, D.C. 20555

Re: LSS Concept/Design Alternatives Summary

Dear Jerry:

We have reviewed the materials provided at the LSS Technical Working Group meeting of January 8, 1992, and have the following comments.

General

I voiced concern during the meeting that we are now postured as examining and recommending for further consideration by Chairman Selin, John Bartlett and the LSS Administrator all technically feasible alternatives to the original SAIC design regardless of ramifications. However, we have not returned to study cost saving alternatives that we initially tabled from earlier consideration because they were so far from the spirit and intent of 10 CFR 2.

At the same time, we all seem to agree that the elements characterized as "LSS Budget Reduction" are not cost saving, but cost reallocations, and an argument could be made that some, in fact, add cost. Most of the items being actively considered merely move the costs to other organizations' budget requests, yet result in essentially the same magnitude costs to be charged against the Nuclear Waste Fund.

Given this, it is dismaying to realize that we have not studied real cost reduction opportunities to the same degree that we have the cost reallocation items. Granted, this is with perfect "20-20 hindsight", and, does not take into account the inherent limitations imposed by the initial rush to meet a November deadline. As a result, after almost four months, we appear to have ignored the opportunity to cost out:

- the run-time CD ROM approach originally broached by OCRWM to OLSSA in July of 1990, or, any other innovative dissemination strategies;
- the true saving that could be achieved if we did not include the full text for every textual document; or,
- the cost impacts of new search and retrieval products that could potentially reduce the need (and very substantial costs) for meeting 99.98% accuracy of ASCII text.

Much mileage has been gotten from the assumption that all of these benefits (reductions/reallocations) can be derived from reuse of InfoSTREAMS technology. Conversely, what if InfoSTREAMS fails? After four months, we have not at all examined any contingency approaches if it does fail.

Comments on the Chart "LSS Concept/Design Alternatives Summary"

Not knowing where this chart may eventually surface without its attendant narrative causes concern over the column titles as stated, the scope of the footnotes, and, the lack of attributions. We feel the following changes should be made:

- the title should read "Summary of OLSSA-Developed LSS Concept/Design Alternatives";
- if there is a column for "LSS Benefit Reduction" there should be one to reflect the attendant "Implementation RISK Increase", which been masked by considering only "Impact on the LSS Rule", (since all the alternatives present additional risk to either LSSA or DOE or NRC or the other parties);
- the column currently entitled "LSS Budget Reduction" should be changed to read "LSS Budget Not in NRC Request to OMB" or even the simpler "Reallocations";
- the column currently entitled "Recommended by Working Group" to be changed to read "Requires Further Evaluation";
- the single asterisk footnote (\*) should be extended at least through the end of FY 2001 to conform to the current schedule for application submission;
- the double asterisk footnote (\*\*) should have a fuller explanation of the assumptions involved, which parties' expectations, etc.;
- there should be an additional footnote that the cost reallocation estimates were developed by OLSSA, that they are estimates derived from yet other estimates provided by SAIC, and, that they are, at best, only "ball-park";
- that there are two elements that are actual cost reductions (10 & 11); and,
- that there are at least three elements where total cost against the Waste Fund could increment even though the reallocation shows a decrease to the NRC/OLSSA budget submittals (2, 7 & 8b).

Finally, the objections I raised about the many "NONE" items under "LSS Benefit Reduction" in my January 2, 1992 letter have not been reflected in any of the materials provided this past Wednesday. I noted in my last letter that

the benefit reduction is not NONE for alternatives #1 and #8; I noted that the reductions are not LITTLE for alternative #7, and indeed it has now become NONE (!); and, I noted that the benefit reduction for alternative #6 was not LITTLE -- that has now been changed to SMALL.

I do not concur with "LSS Benefit Reduction" as it is being presented, and am concerned by the lack of discussion in the narrative. And, I suspect that other parties, as well, will take issue with the subjectivity and presumptive slant of the characterizations. If my comments are not going to be incorporated, I ask that copies of my correspondence at least, then, be included as addenda to the report.

Comments on the Narrative Entitled "LSS Concept/Design Alternatives Summary"

In Item #1, Capture DOE LSS material using INFOSTREAMS versus LSS, add to the last sentence in the last paragraph, ". . . and less the costs necessary for OLSSA to do augmentation." This would imply that the amount (\$76.3M) requires revision. Also in Item #1, the wording in the last paragraph should be changed to reflect a "resultant reallocation" rather than a "resultant savings".

In Item #2, the group noted a change for the chart to wording such as "\$1M or less" and John Voglewede pointed out that it could be either an increment or decrement; the change should be reflected in the wording here. The wording in the last paragraph should be changed to reflect that "Although the cost reallocation is relatively small. . ."

In Item #3, the last paragraph needs a rewrite to reflect that the Working Group is recommending further consideration of the option, and, that it would result in a potential cost reallocation of \$\_\_\_\_. Also, add a paragraph to highlight the inherent difficulties of allocating a user's share of the WAN telecommunication costs.

In Item #4, the last sentence in the first paragraph reads that the "LSS participants have always expected to pay for their own LSS infrastructures, but probably not the workstations." This is ambiguous. What are their infrastructures if not, in fact, the workstations? Also, the last sentence in the last paragraph should be changed to read ". . . the Working Group anticipates that the LSS budget would reflect a \$3.1M reallocation by this alternative."

Item #5 is essentially an internal NRC issue.

Item #6 -- no comments.

For Item #7, I believe that we agreed that this should NOT be recommended for further consideration. Item #7 is, arguably, going to result in an increment to the overall costs against the Waste Fund. Where before there was one high volume printer and LSSA contractor staff to respond to high volume print commands, DOE and NRC must now each separately provide high volume printers and multiple staff, who, by nature of distributing the work, would also be

under-utilized, less-practiced, and subjected to external demands limiting their ability to meet response deadlines.

Item #8a, second paragraph, first sentence should be revised to reflect the group discussion concluding that this should NOT be recommended for further consideration. If Item #8a was going to be recommended, however, it would have to be revisited: OLSSA can not really ever get to a point where they will not require some level of intake processing capability and responsibility. LSSA will have to do record augmentation for the elements I noted in my last letter. LSSA will require capabilities for intake and broadcast of motions practice hardcopy (signed/authenticated) received for entry to the official docket after the electronic versions are submitted. LSSA will need a mechanism for intake of transcripts and setting references to exhibits for materials from the hearings (overnight). Etc., etc.

Item #8b should reflect that this is being recommended for further consideration.

The first sentence of the second paragraph for Item #9 should be revised to read ". . . and provide same to the LSS, reallocating to DOE this share of LSS cost."

Item #10a triggered the objection noted in the opening section: why not reopen the studies of including all documents in full-text versus identifying those sub-groups that would not require/benefit from full text. We agree with the objection made during the group discussion about the logic of recommending an item with MAJOR impact on the but which shows some benefit reduction and less than 1/2% real cost reduction.

Item #10b was noted during the discussion as requiring additional narrative, as it was not intuitively understandable how reducing 75% of the documents held in the system would result in a cost reallocation of only \$5.1 when it implies massive, redundant paper management environments in multiple locations. Again, this is arguably a cost increment to the Waste Fund.

The first sentence in Item #11a should be revised to read "The current OLSSA plan is to provide . . ." The second sentence is wrong. The option does not extend limited access, it delays it for an additional 18 months. Also in Item #11a, the first sentence in the second paragraph should read that "The Working Group recommends further consideration of this alternative . . ."

#### Concerns About Final Report

OCRWM/IMD is willing to provide consensus agreement to recommending all the items for further consideration by the Chairman, the Director, and the Administrator. As I said in the meeting, I will be in the position of having to respond to John Bartlett's request for OCRWM/IMD's analysis of the TWG findings. Even though I am in the position of affirming that they are potential elements, technically feasible, have been examined by the TWG, and should be identified to the recipients of the study, I will be responding to the OCRWM Director's request with pragmatic, DOE-interest, DOE-risk, policy and implementation-based critiques.

For example, for Item #8a, the reality is that there is such great risk to the OCRWM Program, that I could never recommend OCRWM's support. And, I suspect that other potential parties will also have severe problems accepting that DOE would be responsible for processing all of the documents they want introduced to the LSS and getting that material into the LSS prior to 12:00 midnight on the eve of the 6-month lock out prior to the date of the hearings. Will OCRWM be liable for failure to enter the truckloads that could be delivered at "the 11th hour"? How will we prioritize, who coordinate with submitters for load-leveling, would prioritizing another party's documents result in a report of non-compliance by OLSSA that we were not submitting our own materials contemporaneous with their creation? Who is going to pay for incremental resources DOE may have to add at the last minute? And how would DOE procure such last minute resources? InfoSTREAMS is not sized for it. Yet, the TWG would categorize this as having NO LSS Benefit Reduction?

Another example: if we examine Item #8a as a potential cost reallocator, why not just put the same input format and submission requirements on ALL the other parties that DOE is willing to assume for ourselves. The thinking seems to have stopped after reaching the conclusion that "it may as well be DOE's InfoSTREAMS" that shoulders the burdens and responsibilities for everyone else's conversion.

Jerry, I am confident that these concerns will be addressed and expect that this set of materials will make better sense when folded into the report. As I will be in Las Vegas all of next week, I ask that the final draft be FAXed to me in care of John Gandi's office at the Yucca Mt. Project Office. His FAX number is 544-7908 (FTS) or (702)794-7908 (commercial), and his office phone number is 544-7954 (FTS) or (702)794-7954 (commercial) should you need to leave a message for me.

Sincerely,



Daniel J. Graser  
Program Analyst  
Information Management Division  
Office of Civilian Radioactive  
Waste Management

Copies:

B. Cerny, RW-12  
J. Bartlett, RW-1